



OPEN

Gut microbiota of the critically endangered Saiga antelope across two wild populations in a year without mass mortality

Eveliina Hanski^{1,2✉}, Munib Khanyari³, Jingdi Li¹, Kieran A. Bates¹, Steffen Zuther^{4,5}, Martin C. J. Maiden¹, Richard Kock⁶ & Sarah C. L. Knowles^{1✉}

The Saiga are migratory antelopes inhabiting the grasslands of Eurasia. Over the last century, Saiga have been pushed to the brink of extinction by mass mortality events and intense poaching. Yet, despite the high profile of the Saiga as an animal of conservation concern, little is known of its biology. In particular, the gut microbiota of Saiga has not been studied, despite its potential importance in health. Here, we characterise the gut microbiota of Saiga from two geographically distinct populations in Kazakhstan and compare it with that of other antelope species. We identified a consistent gut microbial diversity and composition among individuals and across two Saiga populations during a year without die-offs, with over 85% of bacterial genera being common to both populations despite vast geographic separation. We further show that the Saiga gut microbiota resembled that of five other antelopes. The putative causative agent of Saiga mass die-offs, *Pasteurella multocida*, was not detected in the Saiga microbiota. Our findings provide the first description of the Saiga gut microbiota, generating a baseline for future work investigating the microbiota's role in health and mass die-offs, and supporting the conservation of this critically endangered species.

The Saiga antelope (*Saiga tatarica* ssp.) is a long-distance migratory ungulate, famous for its distinctive pendulant nose (Fig. 1A). Once ranging across nearly the whole of Eurasia, the majority of Saiga now inhabit Kazakhstan, the largest landlocked country in the world. The Kazakh Saiga (*Saiga tatarica tatarica*) is of particular interest due to its vulnerability to mass mortality events, seemingly caused by *Pasteurella multocida*¹. These mass mortality events, together with anthropogenic impacts such as poaching, have reduced the population of Kazakh Saiga by 95% in recent decades, from an estimated 1,200,000 to circa 50,000 between mid-1970s and 2003^{2,3}. Since 2002, the Saiga has been listed as 'critically endangered' by the International Union for Conservation of Nature³; however, many aspects about the biology of the Saiga remain unknown. They live out of sight in the remote grasslands and deserts of Kazakhstan and have proved challenging to keep in captivity⁴, making this species particularly difficult to study. Since around the middle of the twentieth century, the Kazakh Saiga have been living as three geographically distinct populations in the Betpak-Dala, Ustyurt and Ural regions. Both the Betpak-Dala and Ural Saiga populations have suffered from mass die-offs, while residing a minimum of 500 km apart from each other (Fig. 1B).

A better understanding of the biology of the Saiga antelope could provide insights into their susceptibility to mass mortality events with potential conservation application. One key factor that can have important effects on mammal biology is the gut microbiota, the diverse community of microorganisms residing in the intestinal tract. While this community contains microorganisms that range from beneficial to pathogenic, as a whole it provides important functions for the host by regulating key physiological processes such as immune maturation and nutrient extraction^{5,6}. Correspondingly, disruption of the gut microbiota can have adverse effects on the host and increase susceptibility to both infectious and non-infectious diseases⁷.

Many natural factors can influence gut microbiota composition in mammals, including diet, social interactions, infection, and aging^{8–11}. A growing literature indicates how the microbiota of wild animals can be altered by

¹Department of Biology, University of Oxford, Oxford, UK. ²Faculty of Medicine, University of Helsinki, Helsinki, Finland. ³Nature Conservation Foundation, Mysore, India. ⁴Association for the Conservation of Biodiversity of Kazakhstan, Astana, Kazakhstan. ⁵Frankfurt Zoological Society, Frankfurt, Germany. ⁶Centre for Emerging, Endemic and Exotic Diseases, The Royal Veterinary College, University of London, London, UK. ✉email: eveliina.hanski@helsinki.fi; sarah.knowles@biology.ox.ac.uk

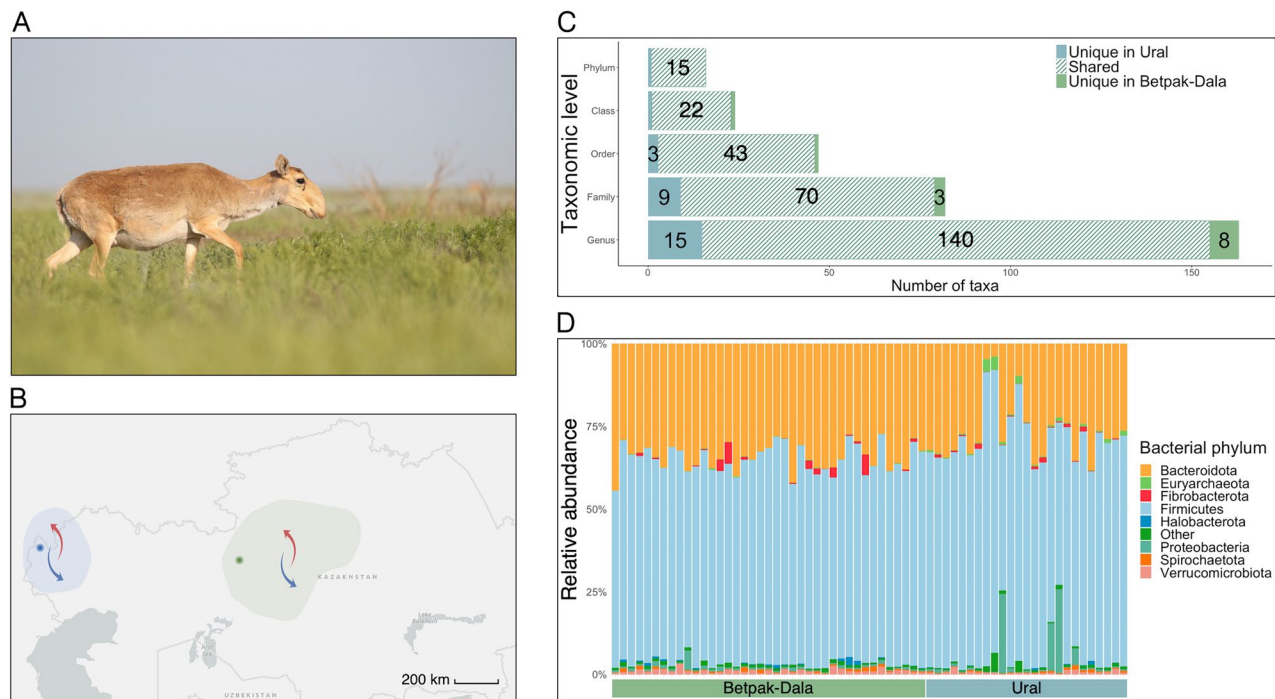


Figure 1. (A) A female Saiga antelope (*Saiga tatarica*). Photo credit: Albert Salemgareyev. (B) Geographic ranges of the two sampled Saiga antelope populations (green, Betspak-Dala population; blue, Ural population). Overall migration directions are indicated with arrows (red, migration direction in spring; blue, migration direction in autumn). Approximate centroids of faecal sample collection sites are indicated with dark dots within the habitats. (C) Number of shared and unique taxa between the two populations at five taxonomic levels. Numbers on bars indicate how many taxa were unique to Ural (blue), Betspak-Dala population (green), or that were shared between the two populations (striped teal). Number of unique/shared taxa is written when ≥ 3 . (D) Gut microbiota composition of the Saiga antelope at phylum level. Rare taxa (mean relative abundance $< 0.03\%$ and prevalence $< 0.10\%$ across samples) are under ‘Other’. Stacked bars are individual samples. Horizontal bars indicate Saiga population.

changes to their environment such as exposure to chemicals, habitat destruction, infectious disease, urbanization, and housing in captivity^{12–20}, yet knowledge about the implications for animal health remains limited. Given the lability and health impacts of the gut microbiota, an improved understanding of host–microbiota interactions also holds conservation potential^{21,22}. For example, characterisation of the Saiga gut microbiota during a year without mass mortality will allow future investigation of the potential role of the microbiota in disease, and might inform the design of microbiota manipulation in captivity, e.g., through diet or faecal microbiota transplants, to facilitate captive breeding programmes for species recovery²¹.

Here, we provide the first characterisation of the Saiga gut microbiota during a year without die-offs, explore the extent of microbiota variation across two geographically distinct populations, and examine how the Saiga microbiota compares with that of other antelope species. These data provide a baseline understanding of the gut microbiota in this critically endangered species, which future work examining the potential significance of the microbiota for mass mortality events can build upon.

Results

Sequencing outcome of the Saiga antelope samples

A total of 80 faecal DNA samples from 70 Saiga antelope (Kazakhstan) (Fig. 1A, B) were sequenced using the Illumina MiSeq platform. Eight Saiga samples were lost during read depth filtering. The remaining 72 samples included eight duplicate samples (duplicate samples from a single faecal deposition, see “Methods”) that were sequenced to assess the robustness of our pipeline. Duplicates generally clustered together with their corresponding samples (Suppl. Fig. 1) and sample ID strongly predicted microbiota composition across these repeat samples, explaining 67.1% of all gut microbial variation (PERMANOVA based on Aitchison distance, $p < 0.001$), indicating the sampling method produced a repeatable representation of the faecal microbiota. These duplicates were removed from the dataset before further analysis. After filtering out singleton and doubleton amplicon sequence variants (ASVs) in the dataset, the two populations of Saiga collectively harboured a total of 4,036 unique ASVs, with a mean of 743 ASVs per sample (range 470–1057). Using the SILVA database (v.138.1), 92.9% ASVs could be assigned to family level, 74.7% to genus level, and 0.7% to species level. Due to the low level of assignment to species level, the highest taxonomic level we considered in analyses was genus.

Composition and diversity of the Saiga antelope gut microbiota

Of the 4,036 ASVs found in Saiga, 2,689 (66.6%) were detected in both populations, while 570 (14.1%) were unique to the Ural population ($n=25$) and 777 (19.3%) were unique to the Betpak-Dala population ($n=39$). Across both populations, 16 bacterial phyla, 24 classes, 47 orders, 82 families, and 163 genera were identified (Fig. 1C). The two populations shared 15 out of 16 phyla (93.8%) and 91.5–91.7% of classes and orders. When inspecting gut microbiota profiles at the family and genus level, 85.4% and 85.9% of taxa were shared between the two populations, respectively. The shared bacterial families and genera formed 93.4–96.1% and 80.3–84.4%, respectively, of the total relative abundance in the Betpak-Dala and Ural populations, such that most bacteria in the Saiga microbiota belong to genera that are common to both populations. Unique genera in each population were not of any particular lower resolution taxa (such as family or class); instead, the Betpak-Dala population had 8 unique genera from 5 different classes. Similarly, the Ural population had 15 unique genera from 8 different classes.

At the phylum level, the Saiga gut microbiota was heavily dominated by Firmicutes and Bacteroidota, which together comprised 94.9% of all reads on average per sample (Fig. 1D). The Ural population harboured one unique phylum, Campilobacterota, that was rare and formed just 0.00002% of total abundance in this population. The shared microbiota (comprised of ASVs found in both populations, irrespective of relative abundance and prevalence) was proportionately dominated by the bacterial families Oscillospiraceae (20.8% and 16.6% of total relative abundance of taxa in Betpak-Dala and Ural, respectively) and Rikenellaceae (15.6% and 9.9%), followed by 68 additional shared families (Suppl. Fig. 2). The shared microbiota contained 140 genera, with the predominant genus being Oscillospiraceae UCG-005 (16.9% and 11.8%). A small number of individuals across the two populations had an increased relative abundance of Proteobacteria (Fig. 1D), primarily driven by *Escherichia/Shigella* as revealed in the genus-level inspection of the Saiga gut microbiota (Suppl. Fig. 2).

Gammaproteobacteria, the class containing *P. multocida*, the bacterium implicated in Saiga mass mortalities¹, was detected in both populations and formed 0.3% (Betpak-Dala) and 2.9% (Ural) of total abundance at class level; however, no ASVs were assigned to the family Pasteurellaceae nor the genus *Pasteurella*. Furthermore, a Nucleotide BLAST search was conducted for ASVs for which SILVA had either assigned the class Gammaproteobacteria or had failed to assign a class. This search did not produce a $\geq 98\%$ species identity match for *P. multocida*, further confirming this bacterium was not detected in this Saiga gut microbiota dataset collected in a year without mass mortalities.

Despite the high proportion of shared ASVs across the two populations (shared ASVs formed 77.6% and 82.5% of all ASVs detected in the Betpak-Dala and Ural populations, respectively), the gut microbiota compositions of the two populations were detectably different, with 10.4% of variation in Aitchison dissimilarity being explained by population identity (PERMANOVA, $p < 0.001$; population identity explained 8.6% of Jaccard dissimilarity, $p < 0.001$). Samples clustered by population when ordinated with the exception of 3 samples, and these patterns were consistent across categorical (Aitchison, Jaccard) and phylogenetic distances (weighted and unweighted UniFrac) (Fig. 2A, Suppl. Fig. 3), indicating consistently that the two populations have distinct gut microbial communities. The three samples that did not cluster by populations were not the samples with increased relative abundance of Proteobacteria (Fig. 1D).

To investigate whether the relative abundance of any specific taxa significantly differed between the two populations, we performed a differential abundance analysis across taxonomic levels from phylum to genus. A total of 14 taxa including three orders, five families, and six genera significantly differed in relative abundance between populations (Fig. 2B). In addition to the shared taxa with significantly different abundances, the Ural population had 9 unique bacterial families including Campylobacteraceae, Moraxellaceae, Acetobacteraceae, Bacteroidales, Micrococcaceae, Peptostreptococcaceae, Coriobacteriales incertae sedis, M2PB4-65 termite group (order Bacteroidales) and p-2534-18B5 gut group (order Bacteroidales). The Betpak-Dala population had 3 unique families; Methanosarcinaceae, Oxalobacteraceae, and Microbacteriaceae. At genus level, the Ural population had 15 unique taxa; *Campylobacter*, *Acinetobacter*, *Acetobacter*, M2PB4-65 termite group incertae sedis, *Alloprevotella*, p-2534-18B5 gut group incertae sedis, Bacteroidales, *Arthrobacter*, *Rummeliibacillus*, *Lactococcus*, *Eubacterium*, *Atopobium*, *Slackia*, *Romboutsia*, and FD2005 (family Lachnospiraceae). Betpak-Dala population had 9 unique taxa; *Methanimicrococcus*, *Rikenella*, GWE2-31-10 (family Spirochaetaceae), *Rathayibacter*, *Pseudoramibacter*, *Bacteroides pectinophilus* group (family Lachnospiraceae), *Oribacterium*, 28-4 (family Lachnospiraceae).

The two Saiga populations had very similar ASV richness (Wilcoxon rank sum test, $p = 0.978$; estimated ASV richness range in Ural Saiga 475–994, mean 759, median 755; range in Betpak-Dala Saiga 498–1087, mean 763, median 761). Similarly, while the Ural Saiga had a slightly higher Shannon diversity index, this difference was not statistically significant (Wilcoxon rank sum test, $p = 0.262$; range in Ural Saiga 113–397, mean 245, median 257; range in Betpak-Dala Saiga 116–325, mean 227, median 232; Fig. 2C).

Comparison between Saiga antelope and five other antelope species

The Saiga gut microbiota was compared with that of five other antelope species for which publicly available 16S rRNA microbiota data could be retrieved: the Tibetan antelope (*Patholops hodgsonii*)²³, Przewalski's gazelle (*Procapra przewalskii*)²⁴, Impala (*Aepyceros melampus*)¹⁸, Springbok (*Antidorcas masupialis*)¹⁸, and Sable antelope (*Hippotragus niger*)¹⁸. The samples originated from China, South Africa, and Namibia (Suppl. Table 1). The investigated antelope species typically live in similar habitats, namely deserts, grasslands, and shrublands³. The target region of 16S rRNA gene varied across the datasets, thus all 16S amplicon sequences were trimmed to the V4 region for this comparative analysis to make them comparable (see “Methods”).

The gut microbiota composition of the Saiga was, at a high taxonomic level, broadly similar to that of the five other antelope species. The predominant phyla in all species were Firmicutes (54.1–70.9% of total abundance) and Bacteroidota (23.8–39.9%). At family level, Oscillospiraceae and Rikenellaceae formed the two predominant

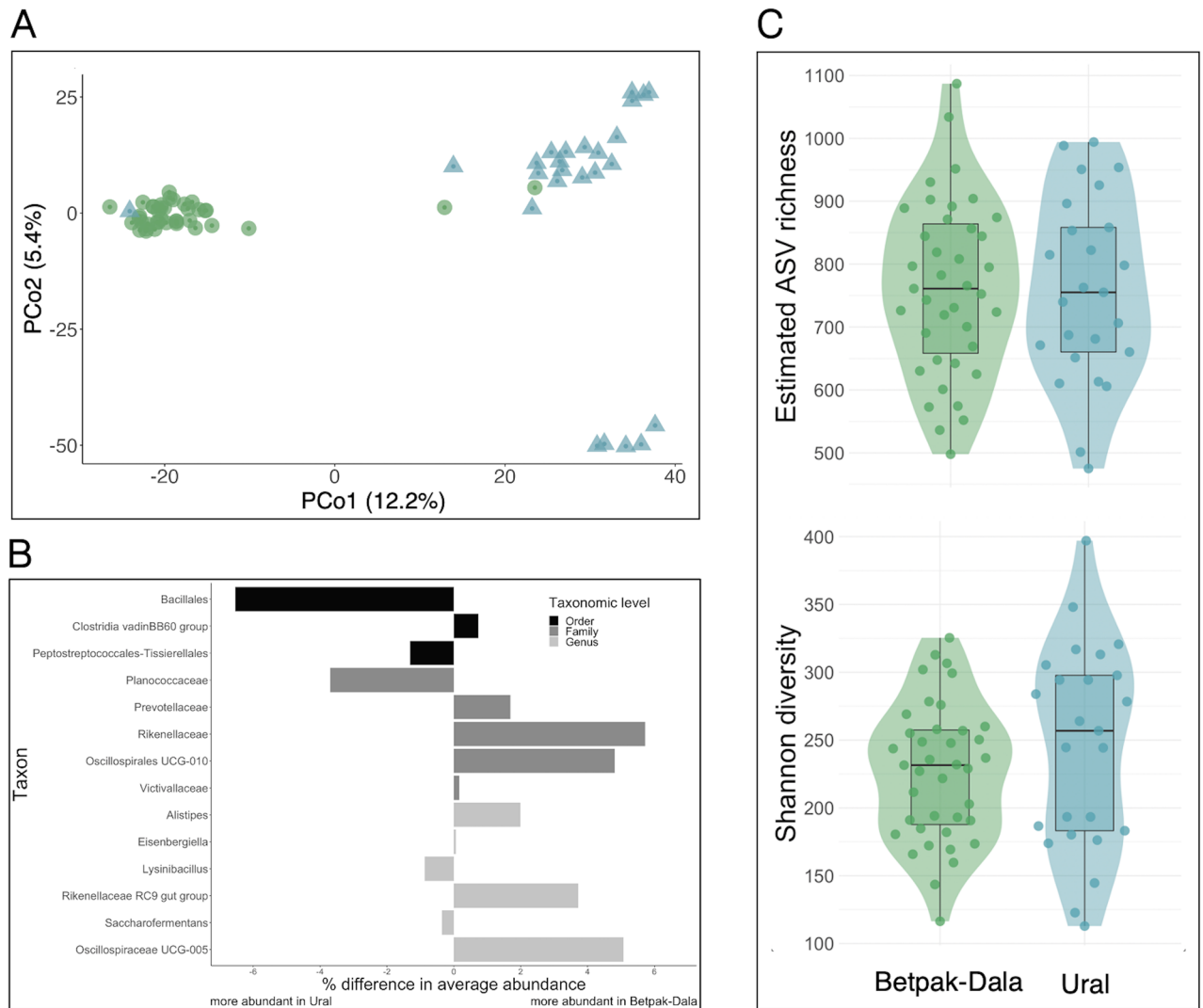


Figure 2. (A) Principal coordinate analysis of Betpak-Dala (green circles) and Ural (blue triangles) Saiga gut microbiota (dis)similarity based on Aitchison distance. Circles and triangles are individual samples. Microbiota similarity increases with proximity between sample points. (B) Bacterial taxa with significantly differing relative abundance between the Betpak-Dala and the Ural Saiga populations. Bars indicate the mean relative abundance in the Betpak-Dala population minus the mean relative abundance in the Ural population. Only taxa with a Benjamini–Hochberg corrected p -value of less than 0.05 and an effect size (standardized mean difference) greater than 1 are shown. (C) Sample-level asymptotic ASV richness (top) and Shannon diversity (bottom) of Betpak-Dala (green) and Ural (blue) Saiga. Circles are individual samples. Horizontal bar indicates median alpha diversity. Statistical differences between Betpak-Dala and Ural Saiga were tested with Wilcoxon rank sum tests ($p > 0.05$ for both ASV richness and Shannon diversity).

bacterial families (25.6–34.2%) in all antelope species except the Tibetan antelope, which had Oscillospiraceae and Lachnospiraceae as the most abundant families, followed by Rikenellaceae (Fig. 3A). At the ASV level, Saiga shared 13.6% and 14.2% of their ASVs with Tibetan antelope and Przewalski's gazelle, respectively (Suppl. Fig. 4A), whereas no common ASVs were detected between Saiga (or the other two Asian antelopes, Przewalski's gazelle and Tibetan antelope) and the African antelopes (Impala, Springbok, and Sable antelope).

Host species identity, which could not be distinguished from dataset identity, had a significant effect on gut microbiota composition explaining 24.1% of variation in bacterial community structure across the dataset (PERMANOVA based on Aitchison distance, $p < 0.001$; host species identity explained 19.5% of Jaccard dissimilarity, $p < 0.001$), although beta dispersion varied significantly among the datasets and could have affected these results ($F = 458.74$, $p = 0.001$). In ordination analyses, Saiga samples clustered separately from samples from all other antelopes on Jaccard, unweighted UniFrac and Aitchison distances but together with other samples on weighted UniFrac distance (Fig. 3B, Suppl. Fig. 4B–D).

The Saiga shared the majority of its gut microbial taxa with at least one antelope at all investigated taxonomic levels (phylum → family). It shared all of its 14 phyla with at least one other antelope species and 12 phyla (85.7%) with at least three other antelope species (Suppl. Fig. 5A). The other antelopes collectively harboured

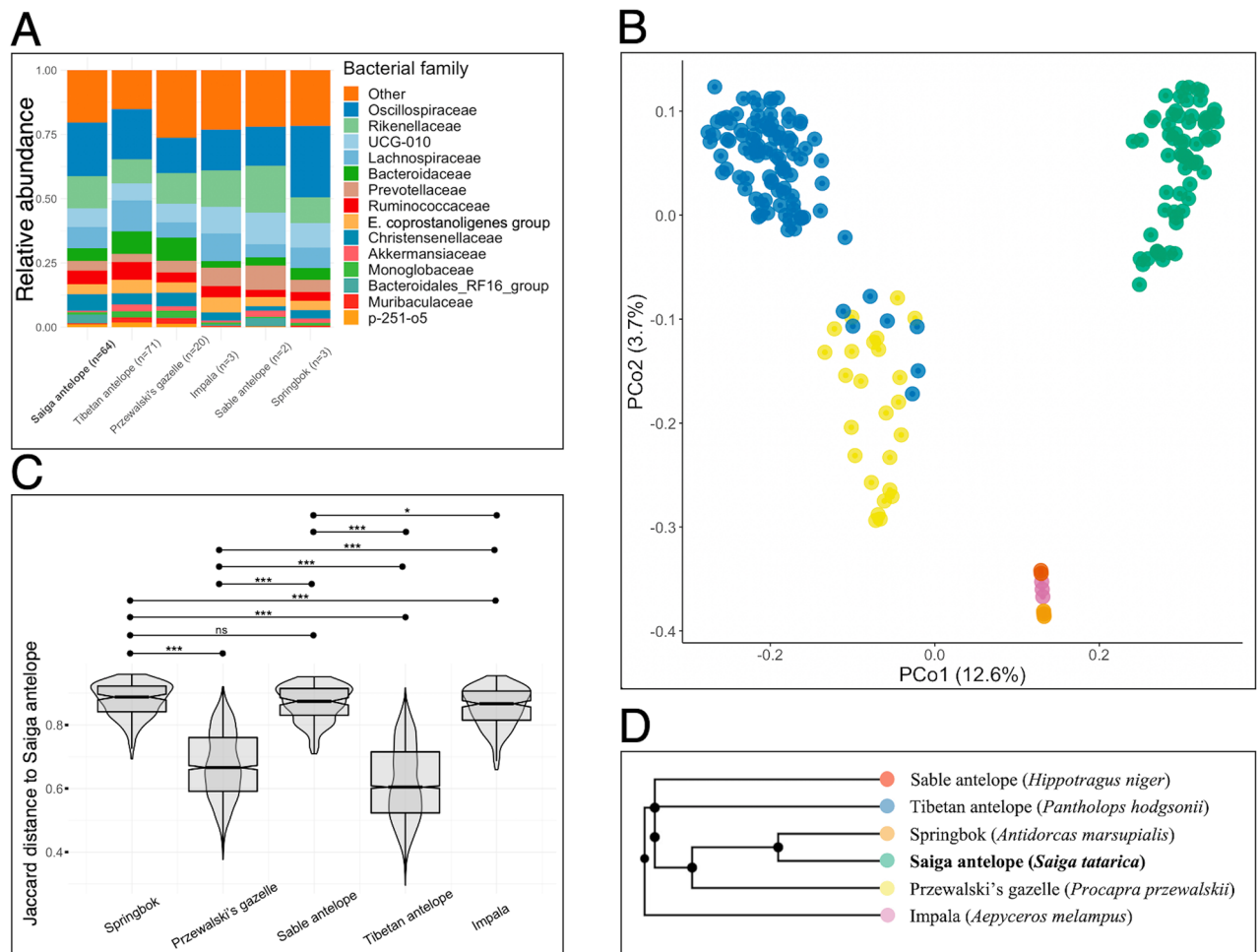


Figure 3. (A) The gut microbiota composition of six antelope species at family level. Rare taxa (mean relative abundance < 0.03% and prevalence < 0.10% across samples) and taxa for which bacterial family could not be assigned are under 'Other'. (B) Principal coordinate analysis of gut microbiota (dis)similarity of six antelope species based on Jaccard distance. Taxa have been agglomerated to the family level before ordination. Circles are individual samples. Colour indicates host species (green, Saiga antelope; yellow, Przewalski's gazelle; blue, Tibetan antelope; orange, Springbok; pink, Impala; red, Sable antelope). (C) Pairwise dissimilarity of bacterial genera (Jaccard distance) between Saiga antelope and (left to right) Springbok, Przewalski's gazelle, Sable antelope, Tibetan antelope, or Impala (antelopes ordered by phylogenetic relatedness to Saiga antelope starting with closest relative). Differences were tested with permutational Wilcoxon rank sum tests (***, $p < 0.001$; *, $p = 0.042$; ns, $p > 0.05$). (D) A cladogram of six antelopes. The nodes indicate a common ancestor and the lines are relative to evolutionary timescale. The cladogram of the six antelopes was retrieved from the TimeTree database.

an additional eight phyla not found in the Saiga (Patescibacteria, Campilobacterota, Myxococcota, Nitrospirota, Acidobacteriota, Chloroflexi, Fusobacteriota, and Synergistota). At class and order levels, the Saiga shared 95.5% and 81.8% of its taxa, respectively, with at least three other antelope species. At family level, the Saiga shared 75.7% of taxa with at least three other antelopes (Suppl. Fig. 5B). These 56 bacterial families formed 94.2% of the total relative abundance of the Saiga gut microbiota. The Saiga harboured two unique bacterial families not detected in any of the other antelope species; Methanosarcinaceae (only detected in Betpak-Dala Saiga) and p-2534-18B5 gut group (order Bacteroidales; only detected in Ural Saiga). These families detected only in Saiga were rare, representing less than 1% of total relative abundance in both Saiga populations.

In terms of shared genera (Jaccard dissimilarity measured at genus-level, due to lack of shared ASVs between Asian and African antelopes), the Saiga microbiota was more similar in composition to that of other Asian antelope species (Tibetan antelope, Przewalski's gazelle) than African species (Sable antelope, Springbok, Impala) (Fig. 3C, D). However, when considering the phylogenetic distance between genera (unweighted UniFrac distance on genera), the Saiga microbiota was more similar to that of African, rather than Asian, antelope species (Suppl. Fig. 6).

Compared to the other antelope species examined, Saiga had the highest alpha diversity (ASV richness and Shannon diversity; Suppl. Fig. 7). The difference in Shannon diversity was significant between Saiga and all other

antelopes (Wilcoxon rank sum tests, $p < 0.02$ for all). However, the only significant difference in ASV richness was between Saiga and Tibetan antelope (Wilcoxon rank sum test, $p < 0.001$; $p > 0.05$ for all other comparisons).

Discussion

The critically endangered Saiga antelope is a species of particular interest because of its vulnerability to mass mortality events. Together with anthropogenic factors, such as poaching and agricultural use, and linear infrastructure expansion, mass mortality events in Saiga can lead to the extinction of regional populations with observed losses of 75% and 88% of the affected population in 1988 and 2015, respectively²⁵. A better understanding of the baseline biology of the Saiga could provide insights into their susceptibility to mass mortality events and tools for more successful husbandry in captivity^{21,22}. As a step towards this, we present a gut microbiota profile of the Saiga, the first to our knowledge, from samples taken in a year without mass mortality and compare it to the gut microbiota of five other antelope species.

Similar to other mammals²⁶, the gut microbiota of the Saiga was dominated by Firmicutes and Bacteroidota, which together formed 95% of total relative abundance. Two subpopulations of the Kazakh Saiga were sampled in the Betpak-Dala and Ural regions. These sampling locations were over 1000 km apart and the estimated geographic ranges of these migrating populations are at least 500 km apart²⁷. Despite the geographic separation, the two populations shared approximately 85–94% of taxa at all inspected taxonomic levels from phylum to genus, and the shared genera formed over 80% of total relative abundance in both populations.

The Betpak-Dala and Ural Saiga populations displayed similar gut microbial alpha diversity, but differed in microbiota composition. Others investigating gut microbiota diversity across two wild populations of the Przewalski's gazelle, a close relative of the Saiga, did not detect significant population differences in either alpha or beta diversity²⁴. These populations are, however, geographically closer to each other (sampling locations < 200 km apart) than the Betpak-Dala and Ural Saiga populations (sampling locations > 1000 km apart), which could contribute to the greater cross-population similarity in the gut microbiota of Przewalski's gazelle compared to Saiga. Geographical proximity could affect gut microbiota similarity through, for instance, similar vegetation type and thus diet as well as population mixing and gut microbe transmission^{9,28}.

Due to the opportunistic nature of sampling and the tendency of the Saiga to avoid people, we were unable to collect more information concerning the individual animals. Hence, it remains unknown how much of the variation within and across the populations was driven by factors such as age and sex. Similarly, while various steps were taken to minimise the potential effect of exposure to oxygen and environmental microbes (see “Methods”), it is possible that the microbial composition of faecal matter was altered by these factors. Considering the two populations inhabit different areas at least 500 km apart, it is likely some of the gut microbial variation was driven by differences in habitat and diet, as has been found in other wild mammalian species²⁹. The faecal samples used in the study were collected during or shortly after the Saiga calving period in May. Hence, the gut microbiota profiles reflect those in the calving period. While this might vary from the Saiga gut microbiota outside calving period, for example due to differences in hormone levels³⁰, this is a relevant time point for sampling the Saiga since the previous mass mortality events have occurred during calving¹, when the birthing females are likely to be stressed and potentially immunocompromised.

As a commensal, *Pasteurella multocida* (the putative causative agent of the Saiga mass mortalities) is most often found in the oral, nasopharyngeal, and respiratory microbiota³¹ and previous work has demonstrated the presence of the bacterium in the respiratory tract of healthy Saiga³². Analysis of tissue samples from the 2015 Saiga mass mortality event provided indication of a possible *P. multocida* invasion from the gut (the bacterium was detected in intestinal mucosa, alongside other body sites¹), thus we searched for the presence of this bacterium in our dataset. We did not detect *P. multocida* in the Saiga gut microbiota and the closest taxonomic rank was the class Gammaproteobacteria.

The absence of *P. multocida* could be an artefact of the method since we characterised the gut microbiota by targeting a region of the 16S rRNA gene, a commonly used method that provides limited fine-scale taxonomic resolution, particularly at species level³³. The method should provide a representative profiling at lower levels, such as at class and family levels³⁴. Considering the class Gammaproteobacteria was the closest taxonomic rank detected to *P. multocida*, our 16S data does suggest the bacterium was not present in the Saiga gut microbiota in this data collected in a year without mass mortalities, or alternatively its abundance was below the detection threshold of the methods used.

To put the Saiga gut microbiota into a wider phylogenetic perspective, we compared it to that of five other antelope species (Tibetan antelope, Przewalski's antelope, Sable antelope, Springbok, and Impala) for which 16S rRNA V4 gut microbiota data from wild individuals was publicly available. Saiga shared the majority of its taxa with other antelope species across taxonomic levels from phylum to genus. Still, microbiota compositions varied significantly between Saiga and the other antelopes. This could have been affected by environmental factors, such as geographical location and diet, as well as experimental factors, such as DNA extraction kit and sequencing batch, which varied between the datasets and could not be controlled for due to small sample size^{28,35,36}. The finding that the hosts shared a high number of taxa at the various taxonomic levels despite these methodological limitations suggests the Saiga does not present an outlier within antelopes from the gut microbiota perspective.

In terms of shared taxa, the Saiga gut microbiota was more similar in composition to antelopes from the same rather than different continent when similarity was measured with non-phylogenetically informed Jaccard distance. However, this conclusion depended on the distance metric considered; when a phylogenetic distance metric was used (unweighted UniFrac), the Saiga microbiota was most similar to that of an African antelope species and its closest relative (Springbok), rather than geographically more proximate Asian antelope species. These results suggest that phyllosymbiosis, under which microbiota similarity is expected to correlate with phylogenetic relatedness and which is observed in several^{37–40} but perhaps not all animals⁴¹, may be detected in the

studied antelope species only when phylogenetic distance of gut microbes is considered. When phylogenetic relatedness of microbes was not considered, geography rather than host phylogeny appeared to have a stronger influence on Saiga microbiota in relation to other antelopes.

Overall, our results indicate the gut microbiota of two geographically disparate Saiga antelope populations is taxonomically rather consistent, but varies in relative abundance of bacterial taxa. We did not detect *Pasteurella multocida*—the bacterium thought to cause Saiga mass mortalities—in the Saiga gut microbiota during this year without die-offs. Finally, we showed that the Saiga gut microbiota resembles that of other antelopes. With this, we provide a baseline description of the gut microbiota in this critically endangered species, on which future work examining the potential role of the microbiota in mass mortality events can build.

Methods

Sample collection

Faecal samples were used to provide a non-invasive characterisation of the gut microbiota. Faecal samples of Saiga antelope were collected opportunistically from two populations in north-west (within 25 kms from 49°59' N, 47°40' E; 'Ural population') and central (within 25 kms from 49°30' N, 61°51' E; 'Betpak-Dala population') Kazakhstan in May 2019, which is during and immediately after the peak calving period.

Faecal samples were collected in a non-intrusive manner. Saiga individuals were located and observed (either directly or through binoculars) until defecation and once they moved on from that location, a fresh faecal sample was collected. The Saiga is the only wild ungulate in the region, with distinct pellet-shaped faeces, and tends to avoid livestock⁴², minimising the risk of faeces mis-identification. During sampling, faecal matter was picked up using clean disposable gloves and dissected with sterile forceps in order to get two separate 100 mg aliquots from inside faecal pellets, to avoid environmental contamination. The time between defecation and collection of samples was within an hour in most cases. While it was not possible to ensure that every faecal sample was from a different individual as Saiga were not marked, our collection methodology made it unlikely to sample the same individual twice: faecal sampling was conducted at various locations over several days within the two sites (Ural, Betpak-Dala) where tens of thousands of Saigas had gathered to calve.

Faecal samples were immediately preserved in DNA/RNA Shield, a preservation solution that protects DNA against degradation and allows sample storage at ambient temperature (Zymo Research, Irvine, California, USA). After a maximum of 3 weeks at ambient temperature in the field, samples were kept at – 20 °C in Kazakhstan until shipping at ambient temperature to the UK in March 2020 (under an import permit from the UK Plant and Animal Health Agency). Upon receipt samples were kept at – 80 °C until DNA extraction.

DNA extraction, PCR amplification, and amplicon sequencing

Genomic DNA was extracted using the ZymoBIOMICS DNA Miniprep Kit, following the manufacturer's protocol (Zymo Research, Irvine, California, USA). A total of 80 samples (32 samples from the Ural population and 48 samples from the Betpak-Dala population) were randomised into four extraction batches. The 80 samples included 10 duplicate aliquots which were used to check the robustness of our sample and data processing pipeline. DNA was eluted in 50 µL DNase-free H₂O, and one negative extraction control (DNase-free H₂O) was included in each extraction batch. Library preparation and sequencing was conducted at the Integrated Microbiome Resource, Dalhousie University, as described in Comeau et al.⁴³. The V4–V5 region of the 16S rRNA gene was amplified by PCR using the 515F–926R primers^{44,45}. All samples were amplified and sequenced in one batch using the Illumina MiSeq platform (Reagent kit v3, 2 × 300 bp chemistry). The sequencing run included a negative control for the PCR reaction and a negative control for the sequencing. All four extraction controls were sequenced either on the sequencing run in question or on subsequent sequencing runs.

Online data acquisition

A Web of Knowledge search was conducted in November 2021 to identify datasets which would allow comparison of the Saiga gut microbiota to that of other wild antelope species. Search keywords included *gut microbiome*, *gut microbiota*, *antelope*, and *ungulate*. Studies of captive animals were excluded. Only publicly available datasets for which V3–V4, V4, or V4–V5 16S rRNA primers and the Illumina sequencing platform had been used were considered. The retrieved and included datasets include the following five antelope species: Tibetan antelope (*Patholops hodgsonii*), Przewalski's gazelle (*Procapra przewalskii*), Impala (*Aepyceros melampus*), Springbok (*Antidorcas masupialis*), and Sable antelope (*Hippotragus niger*) (Suppl. Table 1).

16S data processing

To make all datasets fully comparable, publicly available antelope datasets were downloaded and the raw sequencing reads were processed together with the Saiga sequencing reads using a standardized pipeline, as follows. Downloaded files were converted to match a format that was compatible with Quantitative Insights into Microbial Ecology (QIIME2, 2020.11 distribution) (zipped fastq files with a unique name). First, FastQC v0.11.9⁴⁶ and MultiQC v1.12⁴⁷ were used to visualize read quality, before cutadapt v3.4⁴⁸ was used to remove adapters and/or primers where still present. Due to differences in targeted 16S rRNA region(s) (Suppl. Table 1), for comparative analyses across antelope species, sequencing reads were trimmed to include the V4 region (515F–806R) only to make amplicon sequence variants comparable using cutadapt⁴⁸. In analyses of Saiga microbiota samples alone, such trimming was not performed, in order to provide as much resolution as possible.

Trimmed or original reads containing V4 region were then processed as follows. Low-quality reads were filtered using qiime2 quality-filter q-score (default settings, QIIME2 2020.11) before using the Deblur workflow⁴⁹ to denoise sequences into ASVs, as suggested by QIIME2⁵⁰. Within Deblur, trimming length was determined by manually viewing the quality plot for each study. ASV taxonomy was assigned using a classifier trained on

the full-length 16S rRNA gene SILVA v138.1 database⁵¹, and ASVs taxonomically assigned as ‘mitochondria’ or ‘chloroplast’ as well as those not identifiable at the kingdom level were removed. Phylogenetic trees of the remaining ASVs were built using the SEPP qiime2 plugin (<https://github.com/bioshared/q2-fragment-insertion>) with a reference phylogeny (sepp-refs-gg-13-8).

Negative controls for DNA extraction ($n = 4$) and library preparation ($n = 1$) of Saiga antelope samples collectively contained 13 ASVs with a maximum read count for any given ASV per control of 9. The R package decontam⁵² was used to test for potential contaminants in the Saiga dataset, for which negative controls were available. The decontam test was conducted using the ‘prevalence’ (presence/absence) method, which compares each sequence in biological samples to the prevalence in negative extraction and PCR controls. A sequence was considered a contaminant if it reached a probability of 0.1 in the Fisher’s exact test used in decontam. No potential contaminants were identified in the Saiga dataset with this method.

Sample completeness and rarefaction curves were generated with R package iNEXT^{53,54} for all included datasets combined and the read depth threshold (below which samples were excluded from further analysis) was set at 4000, based on where these curves plateaued. Data were not rarefied. Singletons and doubletons (ASVs with a total of either one or two sequences across the dataset) were removed prior to beta (not alpha) diversity analyses to guard against the possible influence of remaining contaminants and sequencing errors. Microbiome profiles from duplicate Saiga samples were inspected using a principal coordinate analysis (PCoA) based on Aitchison distance in package vegan⁵⁵, and the effect of sample ID for microbiota composition was tested using permutational multivariate analysis of variance (PERMANOVA) on Aitchison distance. Prior to ordination on Aitchison distance, a centered log-ratio (clr) transformation was performed using the package microbiome⁵⁶, with relative abundance values of zero replaced with a pseudocount ($\min(\text{relative abundance}/2)$). Duplicates were removed from the dataset before further analysis.

Analyses

Data was analysed and visualised in R (v4.1.2) using packages phyloseq⁵⁷, vegan⁵⁵, microbiome⁵⁶, ALDEx2⁵⁸, UpSetR⁵⁹, and ggplot2⁶⁰. For the inspection and comparison of taxonomic compositions, ASV counts were transformed into relative abundance per sample. We searched for the presence of *Pasteurella multocida* in the Saiga gut microbiota both by investigating taxonomic assignments from the SILVA database as well as by using the NCBI Nucleotide Basic Local Alignment Search Tool (Nucleotide BLAST). For the latter, we BLAST searched all ASVs either assigned to Gammaproteobacteria (the bacterial class containing *P. multocida*) or not assigned a class by SILVA, against the *P. multocida* type sequence (taxid: 747).

Differential abundance testing of bacterial taxa across the two Saiga populations was conducted using package ALDEx2⁵⁸. Monte-Carlo sampling ($\text{mc.samples} = 128$) from a Dirichlet distribution was used to generate a distribution of clr transformed values for all taxa. Welch’s and Wilcoxon rank sum tests were then performed on the clr transformed values. Taxa with a Benjamini–Hochberg corrected p-value (q-value) of less than 0.05 and an effect size greater than 1 were considered to have a significantly different abundance between the two populations. All taxa found in the Saiga data were included in the differential abundance testing, regardless of whether they are unique to one population or not. Asymptotic alpha diversity (ASV richness and Shannon diversity) was estimated using package iNEXT^{53,54}. Differences in alpha diversity and Jaccard dissimilarity among samples from each population were tested using Wilcoxon rank sum tests.

Beta diversity metrics (Jaccard, Aitchison, unweighted UniFrac, weighted UniFrac) were calculated with the package vegan, and PERMANOVA was performed using Aitchison and Jaccard distances. Ordination was conducted with the package vegan using PCoA. A clr transformation was conducted before ordination for Aitchison distance (zero relative abundances replaced with pseudocounts as described in *Data processing*). Ordination plots were produced using the phyloseq package⁵⁷. A cladogram showing the phylogenetic relationships among the six antelope species was retrieved from the TimeTree database.

Ethics declaration

Faecal samples were collected non-invasively with minimal disturbance to animals.

Data availability

The 16S rRNA amplicon sequencing data of the Saiga antelope have been deposited in GenBank under SRA accession: PRJNA982717.

Received: 7 June 2023; Accepted: 7 October 2023

Published online: 11 October 2023

References

- Kock, R. A. *et al.* Saigas on the brink: Multidisciplinary analysis of the factors influencing mass mortality events. *Sci. Adv.* **4**, e2314 (2018).
- Bekenov, A. B., Grachev, Iu. A. & Milner-Gulland, E. J. The ecology and management of the Saiga antelope in Kazakhstan. *Mamm. Rev.* **28**, 1–52 (1998).
- IUCN. The IUCN Red List of Threatened Species. Version 2022-2 (2022, accessed 29 Nov 2022). <https://www.iucnredlist.org>.
- Rdusch, V., Zimmermann, W., Vogel, K.-H., Ladener, H. & Sliwa, A. Saiga antelopes (Saiga tatarica) at Cologne Zoo: Husbandry experiences and observations. *Zeitschr. des Kölner Zoos.* **59**, 135–149 (2016).
- Semova, I. *et al.* Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell Host Microbe* **12**, 277–288 (2012).
- Chung, H. *et al.* Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* **149**, 1578–1593 (2012).
- Clemente, J. C., Ursell, L. K., Parfrey, L. W. & Knight, R. The impact of the gut microbiota on human health: An integrative view. *Cell* **148**, 1258–1270 (2012).

8. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
9. Raulo, A. *et al.* Social networks strongly predict the gut microbiota of wild mice. *ISME J.* **15**, 2601–2613 (2021).
10. Knutie, S. A. Relationships among introduced parasites, host defenses, and gut microbiota of Galapagos birds. *Ecosphere* **9**, e02286 (2018).
11. Langille, M. G. *et al.* Microbial shifts in the aging mouse gut. *Microbiome* **2**, 50 (2014).
12. Kohl, K. D., Cary, T. L., Karasov, W. H. & Dearing, M. D. Larval exposure to polychlorinated biphenyl 126 (PCB-126) causes persistent alteration of the amphibian gut microbiota. *Environ. Toxicol. Chem.* **34**, 1113–1118 (2015).
13. Kakumanu, M. L., Reeves, A. M., Anderson, T. D., Rodrigues, R. R. & Williams, M. A. Honey bee gut microbiome is altered by in-hive pesticide exposures. *Front. Microbiol.* **7**, 1255 (2016).
14. Amato, K. R. *et al.* Habitat degradation impacts black howler monkey (*Alouatta pigra*) gastrointestinal microbiomes. *ISME J.* **7**, 1344–1353 (2013).
15. Dillard, B. A., Chung, A. K., Gunderson, A. R., Campbell-Staton, S. C. & Moeller, A. H. Humanization of wildlife gut microbiota in urban environments. *Elife* **11**, 89 (2022).
16. Teyssier, A. *et al.* Inside the guts of the city: Urban-induced alterations of the gut microbiota in a wild passerine. *Sci. Total Environ.* **612**, 1276–1286 (2018).
17. Gibson, K. M. *et al.* Gut microbiome differences between wild and captive black rhinoceros—implications for rhino health. *Sci. Rep.* **9**, 7570 (2019).
18. McKenzie, V. J. *et al.* The effects of captivity on the mammalian gut microbiome. *Integr. Comp. Biol.* **57**, 690–704 (2017).
19. Alberdi, A., Martin Bideguren, G. & Aizpurua, O. Diversity and compositional changes in the gut microbiota of wild and captive vertebrates: A meta-analysis. *Sci. Rep.* **11**, 22660 (2021).
20. Malukiewicz, J. *et al.* The gut microbiome of exudivorous marmosets in the wild and captivity. *Sci. Rep.* **12**, 5049 (2022).
21. Treveline, B. K., Fontaine, S. S., Hartup, B. K. & Kohl, K. D. Conservation biology needs a microbial renaissance: A call for the consideration of host-associated microbiota in wildlife management practices. *Proc. Biol. Sci.* **286**, 20182448 (2019).
22. Diaz, J. & Reese, A. T. Possibilities and limits for using the gut microbiome to improve captive animal health. *Anim. Microbiome* **3**, 89 (2021).
23. Shi, Y., Miao, Z.-Y., Su, J.-P. & Wasser, S. K. Shift of maternal gut microbiota of tibetan antelope (*Pantholops hodgsonii*) during the periparturition period. *Curr. Microbiol.* **78**, 727–738 (2021).
24. Liu, R., Shi, J., Shultz, S., Guo, D. & Liu, D. Fecal bacterial community of allopatric Przewalski's gazelles and their sympatric relatives. *Front. Microbiol.* **12**, 737042 (2021).
25. Robinson, S. *et al.* Opportunistic bacteria and mass mortality in ungulates: Lessons from an extreme event. *Ecosphere* **10**, e02671 (2019).
26. Rojas, C. A., Ramirez-Barahona, S., Holekamp, K. E. & Theis, K. R. Host phylogeny and host ecology structure the mammalian gut microbiota at different taxonomic scales. *Anim. Microbiome* **3**, 33 (2021).
27. ACBK. In *Association for the Conservation of Biodiversity of Kazakhstan. Annual Report* (2011).
28. Knowles, S. C. L., Eccles, R. M. & Baltrūnaitė, L. Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol. Lett.* **22**, 826–837 (2019).
29. Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R. P. & Pringle, R. M. Covariation of diet and gut microbiome in African megafauna. *Proc. Natl. Acad. Sci.* **116**, 23588–23593 (2019).
30. Mallott, E. K., Borries, C., Koenig, A., Amato, K. R. & Lu, A. Reproductive hormones mediate changes in the gut microbiome during pregnancy and lactation in Phayre's leaf monkeys. *Sci. Rep.* **10**, 9961 (2020).
31. Wilson, B. A. & Ho, M. *Pasteurella multocida*: From zoonosis to cellular microbiology. *Clin. Microbiol. Rev.* **26**, 631–655 (2013).
32. Sanchez-Monge, F. Calving status and commensalism of *Pasteurella multocida* in the surviving Betspak-dala saiga population in May 2016, after a mass mortality event in May 2015. In *Thesis, Royal Veterinary College, University of London* (2016).
33. Johnson, J. S. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**, 5029 (2019).
34. Heidrich, V. *et al.* Choice of 16S ribosomal RNA primers impacts male urinary microbiota profiling. *Front. Cell Infect. Microbiol.* **12**, 862338 (2022).
35. Zhu, L. New challenges in the study of the evolution of wild animals and their gut microbiome. *Ecol. Evol.* **12**, e8904 (2022).
36. Shaffer, J. P. *et al.* A comparison of six DNA extraction protocols for 16S, ITS and shotgun metagenomic sequencing of microbial communities. *Biotechniques* **73**, 34–46 (2022).
37. Song, S. J. *et al.* Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *mBio* **11**, 896 (2020).
38. Kohl, K. D., Dearing, M. D. & Bordenstein, S. R. Microbial communities exhibit host species distinguishability and phyllosymbiosis along the length of the gastrointestinal tract. *Mol. Ecol.* **27**, 1874–1883 (2018).
39. Weinstein, S. B. *et al.* Microbiome stability and structure is governed by host phylogeny over diet and geography in woodrats (*Neotoma* spp.). *Proc. Natl. Acad. Sci.* **118**, e2108787118 (2021).
40. Ochman, H. *et al.* Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol.* **8**, e1000546 (2010).
41. Moeller, A. H. *et al.* Dispersal limitation promotes the diversification of the mammalian gut microbiota. *Proc. Natl. Acad. Sci.* **114**, 13768–13773 (2017).
42. Singh, N. J. & Milner-Gulland, E. J. Conserving a moving target: Planning protection for a migratory species as its distribution changes. *J. Appl. Ecol.* **48**, 35–46 (2011).
43. Comeau, A. M., Douglas, G. M. & Langille, M. G. I. Microbiome helper: A custom and streamlined workflow for microbiome research. *mSystems* **2**, 796 (2017).
44. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).
45. Walters, W. *et al.* Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1**, 1 (2016).
46. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. (2010). Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
47. Ewels, P., Magnusson, M., Lundin, S. & Källér, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
48. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
49. Amir, A. *et al.* Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* **2**, 2 (2017).
50. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
51. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
52. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).

53. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: Interpolation and Extrapolation for Species Diversity. (2022). R package version 3.0.0, http://chao.stat.nthu.edu.tw/wordpress/software_download/.
54. Chao, A. *et al.* Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* **84**, 45–67 (2014).
55. Oksanen, J. *et al.* vegan: Community Ecology Package. R package version 2.6-4. (2022)
56. Lahti, L. & Shetty, S. *microbiome R package*. (2017). <http://microbiome.github.io>.
57. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
58. Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G. & Gloor, G. B. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS One* **8**, e67019 (2013).
59. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
60. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, New York, 2016). ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.

Acknowledgements

EH was supported by a scholarship from the Osk. Huttunen Foundation (Finland). MK was supported by the National Geographic Society's grant (EC-54710C-19). The research was funded by grants from the Emil Aaltonen Foundation (Finland), and fieldwork was supported by the Altyn Dala Conservation Initiative. SCLK was supported by a NERC fellowship (NE.L011867/1) and funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 851550).

Author contributions

E.H. designed the Saiga antelope sample collection protocol. R.K. and M.K. collected the Saiga antelope samples. S.Z. assisted with fieldwork logistics. E.H. conducted DNA extraction. J.L. and E.H. processed the raw data. E.H. analysed the data with input from K.A.B. and S.C.L.K., E.H., R.K., S.C.L.K., and M.C.J.M. contributed to study conceptualisation. E.H. wrote the manuscript with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44393-z>.

Correspondence and requests for materials should be addressed to E.H. or S.C.L.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023